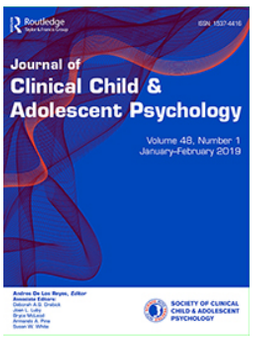

Routledge
Taylor & Francis Group

Journal of Clinical Child & Adolescent Psychology

ISSN: 1537-4416 (Print) 1537-4424 (Online) Journal homepage: https://www.tandfonline.com/loi/hcap20

Core Elements of Family Therapy for Adolescent Behavior Problems: Empirical Distillation of Three Manualized Treatments

Aaron Hogue, Molly Bobek, Sarah Dauber, Craig E. Henderson, Bryce D. McLeod & Michael A. Southam-Gerow

To cite this article: Aaron Hogue, Molly Bobek, Sarah Dauber, Craig E. Henderson, Bryce D. McLeod & Michael A. Southam-Gerow (2019) Core Elements of Family Therapy for Adolescent Behavior Problems: Empirical Distillation of Three Manualized Treatments, Journal of Clinical Child & Adolescent Psychology, 48:1, 29-41, DOI: 10.1080/15374416.2018.1555762

To link to this article: https://doi.org/10.1080/15374416.2018.1555762

Journal of Clinical Child & Adolescent Psychology, 48(1), 29–41, 2019

ISSN: 1537-4416 print/1537-4424 online
DOI: https://doi.org/10.1080/15374416.2018.1555762

Core Elements of Family Therapy for Adolescent Behavior Problems: Empirical Distillation of Three Manualized Treatments

Aaron Hogue, Molly Bobek, and Sarah Dauber
Center on Addiction

Craig E. Henderson
Department of Psychology, Sam Houston State University

Bryce D. McLeod and Michael A. Southam-Gerow
Department of Psychology, Virginia Commonwealth University

Family therapy has the strongest evidence base for treating adolescent conduct and substance use problems, yet there remain substantial barriers to widespread delivery of this approach in community settings. This study aimed to promote the feasibility of implementing family-based interventions in usual care by empirically distilling the core practice elements of three manualized treatments. The study sampled 302 high-fidelity treatment sessions from 196 cases enrolled in 1 of 3 manualized family therapy models: multidimensional family therapy (102 sessions/56 cases), brief strategic family therapy (100 sessions/94 cases), or functional family therapy (100 sessions/46 cases). Adolescents were 57% male; 41% were African American, 31% White non-Hispanic, 9% Hispanic American, 6% another race/ethnicity, and 13% unknown. The observational fidelity measures of all three models were used to code all 302 sessions. Fidelity ratings were analyzed to derive model-shared treatment techniques via exploratory factor analyses on half the sample; the derived factors were then validated via confirmatory factor analyses supplemented by Bayesian structural equation modeling on the remaining half. Factor analyses distilled 4 clinically coherent practice elements with strong internal consistency: Interactional Change (6 treatment techniques; Cronbach's $\alpha = .93$), Relational Reframe (7 techniques; $\alpha = .79$), Adolescent Engagement (4 techniques; $\alpha = .68$), and Relational Emphasis (4 techniques; $\alpha = .67$). The 4 empirically derived factors represent the core elements of 3 manualized family therapy models for adolescent behavior problems, setting the foundation of a more sustainable option for delivering evidence-based family interventions in routine practice settings.

Public Health Significance: Increasing implementation of high-fidelity family-based interventions would improve the quality of treatment services for adolescent conduct and substance use problems.

This study advances efforts to disseminate family-based treatment in routine care for adolescent conduct and substance use problems by identifying core practice elements of the family therapy (FT) approach. Core elements are discrete treatment practices common to multiple treatment

Correspondence should be addressed to Aaron Hogue, Center on Addiction, 633 Third Avenue, 19th Floor, New York, NY 10017. E-mail: ahogue@centeronaddiction.org

manuals for a given behavioral disorder that can be more readily adopted and flexibly implemented than full manuals (Chorpita, Daleiden, & Weisz, 2005). This study distilled the core practices of three well-validated family therapy models via observational coding of high-fidelity treatment sessions.

Formidable Barriers to Implementing Manualized Family Therapy in Routine Care

Family therapy is an evidence-based approach to treating adolescent behavior problems that focuses on intervening directly with family members to repair relationships and addressing challenges encountered by adolescents and caregivers in key extrafamilial systems (Baldwin, Christian, Berkeljon, & Shadish, 2012). There are a handful of "brand name" family therapy models designed to treat adolescent behavior problems (ABPs), with three establishing themselves as efficacious for disruptive behavior, substance use, or both: functional family therapy (FFT), multidimensional family therapy (MDFT), and brief strategic family therapy (BSFT; Hogue, Henderson, Becker, & Knight, 2018; McCart & Sheidow, 2016). Manualized family therapy models for ABPs have posted an exemplary record of success in comparison to individual and group treatments and, in meta-analyses, produced the largest effect sizes by a healthy margin (Baldwin et al., 2012; Riedinger, Pinquart, & Teubert, 2017; Tanner-Smith, Wilson, & Lipsey, 2013).

Despite their exceptional research portfolio, manualized FT models have not been widely adopted in behavioral treatment systems (Riedinger et al., 2017). Developers of manualized FTs typically disseminate their respective models by establishing corporate entities that contract directly with host agencies to govern adoption and training activities. To promote high-fidelity delivery, each brand-name model contains an extensive set of quality assurance procedures anchored by a standardized training tool kit, guidelines for ongoing training and observational consultation from model experts, and implementation support and fidelity tracking methods that feed therapy session data back to providers (see Hogue, Ozechowski, Robbins, & Waldron, 2013). Such procedures incur substantial costs numbering tens of thousands of dollars annually for initial training plus certification maintenance (Substance Abuse and Mental Health Services Administration, 2012). In addition, manualized FTs prescribe numerous complex treatment procedures, often with a fixed intervention sequence—features that inhibit treatment selection and tailoring practices for clients favored by community clinicians (Chorpita et al., 2005).

Core Elements: Innovative Approach to Implementing Evidence-Based Interventions

Implementation barriers such as those just listed are common to manualized treatments of many kinds, not just FT models. In response, experts in mental health (e.g., Chorpita, Becker, & Daleiden, 2007; Garland, Hawley, Brookman-Frazee, & Hurlburt, 2008) and substance use advocate an alternative strategy to complement manual-driven methods: Focus on core elements of evidence-based interventions (EBIs) that represent a reduced set of intervention techniques common to multiple treatments for a given disorder. This is achieved by (a) specifying the discrete techniques prescribed by similar treatment manuals and (b) distilling these techniques into a smaller number of overlapping practice elements that are core features of each manual (Chorpita & Daleiden, 2009). Thus, whereas manuals are predominantly complex, uniform, and disorder specific, distilled core elements are instead granular, flexible, and transdiagnostic—three user-centered features (Lyon & Koerner, 2016) that may help solve vexing barriers associated with implementing manuals in routine care (Weisz, Bearman, Santucci, & Jensen-Doss, 2017). For example, core practices can be more readily learned by clinicians and flexibly applied to a larger set of problems, making them well suited for client-tailored treatment planning (Garland, Bickman, & Chorpita, 2010). The transdiagnostic feature of core practices is highly germane to ABPs, insofar as conduct problems and substance use problems share a large set of common risk and protective factors during adolescence, for which a core set of FT techniques appears applicable and effective (see Baldwin et al., 2012; Henggeler & Schaeffer, 2016; Hogue et al., 2015).

In the most ambitious distillation project yet undertaken, Chorpita and Daleiden (2009) reviewed 322 randomized trials of EBIs for youth populations across the behavioral spectrum. They reliably isolated 41 core elements variously used to treat anxiety and depression, conduct problems, attention and impulsivity, sleep and eating problems, autism-spectrum problems, substance use, and so forth. Notably, from among the numerous trials containing a FT condition, the authors identified only one undifferentiated code broadly termed "family therapy." This underscores the absence of progress to date in distilling core elements for the FT approach.

Family Therapy for Adolescent Behavior Problems: Primed for Core Elements Distillation

The FT approach for ABPs is well positioned to support a distillation process with far more differentiated results than the solitary "family therapy" element isolated in the Chorpita and Daleiden (2009) taxonomy. There is ample evidence from existing research that three brand-name FT models for ABPs—FFT, MDFT, BSFT—share a set of common treatment techniques that are operationalized in their respective model fidelity metrics, are empirically linked to client outcomes, and can be readily synthesized into core elements. FFT contains family interventions such as minimizing blame (Alexander et al., 1989), reframing

negative behaviors (Robbins, Alexander, Newell, & Turner, 1996), and interrupting defensive interactions among family members (Robbins, Alexander, & Turner, 2000) that have been correlated with in-session and posttreatment outcomes. MDFT contains several techniques ubiquitous within the FT approach (e.g., coach in-session family interactions; target caregivers for change) that predict long-term improvement in family functioning as well as adolescent symptoms (Hogue, Dauber, Samuolis, & Liddle, 2006; Hogue et al., 2008). BSFT is anchored by signature techniques of the FT approach—joining, diagnostic enactment, reframing, and restructuring interventions—that variously predict improvements in treatment engagement, family functioning, and adolescent symptoms (Robbins et al., 2011).

Buoyed by this cohesive portfolio of process-outcome research, the authors completed a conceptually driven distillation process to isolate the core elements of these manualized FT models (Hogue et al., 2017). Our conceptual distillation process focused on the observational fidelity scales corresponding to each model: FFT (Ozechowski & Waldron, 2016), MDFT (Hogue et al., 1998), and BSFT (Hervis & Robbins, 2015); see the Study Measures section. This process thereby leveraged the available fidelity blueprints laid bare by the respective model developers to operationalize discrete techniques considered essential to model adherence. In accord with standard distillation procedures (Chorpita & Daleiden, 2009), the various treatment techniques contained in the three fidelity scales were examined to identify thematic clinical strategies that appeared to be (a) common across the models, (b) theoretically salient to the FT approach, and (c) embodied by multiple techniques from all three scales. This process yielded four core practice elements: (a) Family Engagement: enhance family member involvement in treatment and build the relationship between therapist and all members, (b) Relational Reframing: deemphasize individual and intrapsychic ways of defining problems in favor of a systemic conceptualization focused on relational processes, (c) Family Behavior Change: teach new skills and encourage individual behavior changes that promote improved family relations, and (d) Family Restructuring: prompt changes in family emotional processes and encourage insight into predominant cycles of relational interactions.

Study Aim: Identify EBI Core Elements via Novel Empirical Distillation Methods

The aim of the current study was to empirically distill core elements of FT for ABPs via observational ratings of therapy sessions using model-specific fidelity scales. To our knowledge this study is the first to use empirical distillation methods to identify EBI core elements. Prior distillation efforts have used exclusively conceptual methods, which rely on expert review of the content of treatment manuals and/or protocol descriptions, usually fortified by team-based coding and consensus procedures among multiple reviewers (e.g., Chorpita & Daleiden, 2009), and sometimes further confirmed by expert survey (e.g., Garland et al., 2008). Conceptual methods possess the considerable virtues of strong face validity, modest technical demands, and flexibility in review procedures. By the same token, they are subject to legitimate questions regarding the reliability and generalizability of distillation results: How credibly do the distilled elements represent the original EBI content? Would different groups of experts reach meaningfully different results, and if so, what are the implications for the clinical validity of the elements themselves? These questions loom largest for complex EBIs that resist disaggregation or easy reduction to simpler constructs, such as manualized FT (Chorpita & Daleiden, 2009).

To meet our study aim we developed a novel, two-part empirical distillation method. First, we collected observational fidelity data on 302 treatment sessions of FFT, MDFT, and BSFT. As described in the Method section, all sessions were verified to be high-fidelity exemplars of the respective models, thereby constituting a "gold standard" sample pool that incontrovertibly represented the models and their constituent treatment techniques. We employed a fully crossed design in which observers rated each of the 302 sessions for the occurrence of all techniques from all three fidelity scales. Although time-intensive and technically demanding, nonparticipant ratings remain the most rigorous method for assessing treatment fidelity in both research trials (Hogue, Liddle, & Rowe, 1996) and standard practice settings (Garland et al., 2010). Second, we conducted factor analyses of the fully crossed observational data set to distill its latent, core FT elements. We randomly split the sample in half, deriving core factors via exploratory factor analyses and then validating the derived factors via confirmatory factor analyses supplemented by Bayesian analysis. We conservatively hypothesized that these empirical methods would generate core FT elements mirroring those obtained from our conceptual distillation: Family Engagement, Relational Reframing, Family Behavior Change, and Family Restructuring.

METHOD

Study Sample: Three High-Fidelity Sample Pools

This study sampled 302 videotaped sessions from 196 cases enrolled in three manualized FT models for ABPs: MDFT (101 sessions/56 cases), BSFT (100 sessions/94 cases), and FFT (100 sessions/46 cases). Adolescents who attended sessions were 57% male; 41% were African American, 31% White non-Hispanic, 9% Hispanic American, 6% another race/ethnicity, and 13% unknown. MDFT sessions were selected from a controlled prevention trial involving high-risk young adolescents (11–14 years; Hogue, Liddle,

Becker, & Johnson-Leckrone, 2002). Trial therapists demonstrated strong adherence to the MDFT model via observational ratings (Hogue, Liddle, Singer, & Leckrone, 2005). There were five male therapists: two African American, two White non-Hispanic, and one Asian American. BSFT sessions were selected from an archive of expert-supervised training sessions in which community therapists treated adolescents (13–18 years) in outpatient care for ABPs. All BSFT sessions were observationally rated by model experts as demonstrating above-average model fidelity; sessions included 53 therapists from diverse backgrounds (specific race/ethnicity data were not available). FFT sessions were selected from an archive of controlled trials involving adolescents (14–19 years) with conduct and substance use disorders (e.g., Rohde, Waldron, Turner, Brody, & Jorgensen, 2014; Waldron, Slesnick, Brody, & Turner, 2001). Sessions were chosen from among those having been observationally rated as exemplary in model fidelity. There were 11 FFT therapists (nine female); race/ethnicity data were not available. Establishing the strong fidelity of sampled sessions is essential to ensure that distilled techniques validly represent the three manualized models; however, source fidelity data from the three sample pools were not used in study analyses.

Sample Selection Procedures

Across all three sample pools, sessions were randomly chosen to maximize representation across treatment duration: one each from Sessions 1–3 (Early), 4–9 (Middle), and 10+ (Later), depending on availability (as described next). Only sessions lasting at least 30 min were retained. Adolescents and caregivers appeared together in 88% of sessions, with 7% of sessions containing only adolescents and 5% only caregivers. The MDFT pool (102 sessions/56 cases) was randomly selected from 574 recorded sessions of the 56 cases: 28% Early, 34% Middle, and 38% Later. The BSFT pool (100 sessions/94 cases) was chosen by model experts with personal knowledge of each case. Because this was a training sample, almost every available BSFT case had only one recorded high-fidelity session, and the bulk of recorded sessions represented early treatment: 52% Early, 36% Middle, 12% Later. The FFT pool (100 sessions/46 cases) was randomly selected to maintain a balance of cases deemed by model experts to be "easy" versus "difficult": 27% Early, 51% Middle, 22% Later.

Study Measures

Multidimensional Family Therapy Therapist Behavior Rating Scale (MDFT-TBRS)

The 13-item MDFT-TBRS (Hogue et al., 1998) describes treatment techniques for engaging family members in therapy, improving parenting and family skills, intervening in family interaction patterns, and addressing key developmental issues involving extrafamilial systems. The scale measures MDFT extensiveness, that is, the thoroughness/frequency with which each technique was used in the observed session, based on a 5-point Likert-type scale: 1 (not at all), 2 (a little bit), 3 (moderately), 4 (considerably), 5 (extensively). The scale has shown strong factor properties, interrater reliability (ICCs = .60–.89), internal consistency (α = .67–81), and links to client outcomes in studies of high-risk youth (Hogue et al., 2005) and youth with conduct and substance use disorders (Hogue et al., 2006, 2008).

Brief Strategic Family Therapy Fidelity Rating Scale (BSFT-FRS)

The 24-item BSFT-FRS (Hervis & Robbins, 2015) specifies clinical techniques associated with four domains of structural-strategic FT: Joining (alliance building with all family members), Tracking (assessing family relations in session), Reframing (endorsing positive meaning to attributions/behaviors), Restructuring (realigning family boundaries and strengthening relationships). Analyses from a multisite effectiveness study (Robbins et al., 2011) revealed strong interrater reliability across model domains (intraclass correlation coefficient [ICCs] = .81–.85), excellent construct validity (measurement model yielded comparative fit index [CFI] = .94 and root mean square error of approximation [RMSEA] = .08 on cross-validation sample; CFI = .94 and RMSEA = .08 on full sample), and robust item-factor loadings and composite reliability: .32–.77 (ICC = .85) for Joining, .48–.86 (ICC = .78) for Tracking, .99–1.00 (ICC = .99) for Reframing, and .32–.82 (ICC = .83) for Restructuring, with factor intercorrelations ranging from .43-.80. BSFT-FRS fidelity ratings predicted client retention and symptom improvement. For consistency, in the current study the BSFT-FRS used the same 5-point extensiveness anchors described for MDFT-TBRS.

Functional Family Therapy Therapist Adherence Rating Scale (FFT-TARS)

The 14-item FFT-TARS (Ozechowski & Waldron, 2016) describes treatment techniques that capture five basic phases of the FFT model: engagement, motivation, relational assessment, behavior change, generalization. FFT focuses on establishing balanced alliances with family members, assessing the relational functions of individual behaviors, using reframing and relabeling techniques as meaning-changing interventions, and teaching new family skills that generalize to multiple contexts (Alexander, Waldron, Robbins, & Neeb, 2013). FFT-TARS items derive from forerunning FFT observational coding scales that demonstrated strong interrater reliability, internal consistency, construct validity, and predictive validity in numerous process-outcome studies (e.g., Alexander et al., 1989; Robbins et al.,

1996, 2000). Again, in the current study the FFT-TARS used the same 5-point extensiveness anchors just presented.

Observational Coding Procedures and Raters

A separate cohort of observational raters was trained to code each of the three sample pools. Raters in all three cohorts were trained during twice-weekly meetings via review of the relevant rating manual, in-group coding practice, and exercises to increase understanding of scale items. Study coding in each cohort commenced once raters reached a collective reliability threshold of ICC (Shrout & Fleiss, 1979) = .65 for 80% of items, and monitored thereafter. All sessions were independently coded in their entirety by two raters randomly assigned to sessions in pairs according to a balanced incomplete randomized block design (Fleiss, 1981). The MDFT rater cohort contained 14 women (22–-32 years of age): nine White non-Hispanic, two African American, one Hispanic American, one Asian American. Seven had a master's degree, and five were enrolled in a graduate psychology program. The BSFT cohort contained six women (22–32 years): five White non-Hispanic and one African American. Four had a master's degree. The FFT cohort contained six women (23–33 years): four White non-Hispanic, one African American, one Hispanic American. Four had a master's.

Plan of Analysis

Study analyses occurred in four stages. In Stage 1: item selection, interrater reliability and descriptive statistics for each item on each coding scale were calculated separately for each sample pool. Interrater reliability was calculated using the one-way random ICC(1, 2; Shrout & Fleiss, 1979). Based on the following decision rules, we excluded 13 items (two MDFT, six BSFT, five FFT) from further analysis: (a) items with an ICC less than .30 and with a p value above .10 in at least one sample pool and (b) items with skew or kurtosis above 10 in at least one sample pool. The remaining 38 items proceeded to Stage 2 analyses. Prior to Stage 2 analyses, item scores were averaged across both raters to yield a single score for each item within each sample pool. The sample was randomly split into two half samples to allow for exploratory factor analysis (EFA) on one half followed by confirmatory factor analysis (CFA) on the other. The random split was conducted within each of the sample pools to ensure equal representation of each sample pool within each half sample.

In Stage 2: EFA, we used principal components EFA with direct oblimin rotation, applying the sandwich variance estimator to account for the nesting of sessions within therapists. Because our study aim was to describe how the various observed techniques clustered together in this set of exemplary sessions, principal components EFA, which describes all variance in the data set (i.e., how experts delivered each model), is an appropriate approach for distilling core practices. EFA was conducted on the first half sample ($N = 151$), and one through six factor solutions were extracted. The four-factor solution (see the Results section) was ultimately selected based on eigenvalue decline and overall interpretability; items with factor loadings less than .40 were trimmed from the final model solution to maximize parsimony (Tabachnick & Fidell, 2007).

Following EFA, Stage 3: CFA was conducted on the other half sample ($n = 151$) to confirm the fit of the four-factor solution derived from EFA. First, preliminary CFA models were estimated for each of the factors separately to achieve adequate fit for each factor individually, before attempting to fit the full four-factor model. Modification indices were examined to guide model adjustments needed to maximize model fit, and poorly performing items were trimmed as needed. Model fit was assessed using the model chi-square statistic and two supplementary fit indices, RMSEA and CFI. RMSEA values of .06 and below, and CFI above .95, indicate strong model fit; CFI $\geq$ .90 and RMSEA $\leq$ .08 indicate adequate fit (Browne & Cudeck, 1993; McDonald & Ho, 2002). Both EFA and CFA were conducted in Mplus 7.31 (Muthén & Muthén, 1998–2017). The sandwich variance estimator was used to account for the nesting of sessions within therapists (Asparouhov, 2005).

As described in the Results section, the full four-factor CFA model failed to converge. Muthén and Asparouhov (2012) argued that the typical implementation of CFA modeling using maximum-likelihood estimation applies overly strict assumptions that often lead to model misfit, triggering a series of model modifications that may capitalize on chance and fail to replicate in future studies. As an alternative they proposed using Bayesian analysis to relax these assumptions and provide "wiggle room" for researchers to specify small non-zero ranges into which these estimates can fall. Bayesian estimation provides a means for incorporating a priori knowledge from previous research directly into the statistical model (Ozechowski, 2014; van de Schoot et al., 2014). This is accomplished by specifying a prior distribution, which expresses a priori information about how the parameters estimated in the statistical model are distributed in the population (Ozechowski, 2014). A Bayesian analysis involves three elements: (a) prior knowledge on the parameter being tested, captured by the prior distribution (parameter estimate and its associated variance); (b) information provided by the data at hand (likelihood function); and (c) the posterior distribution, which represents the combination of the two previous elements and is derived using Bayes's theorem. The point estimate of the parameter of interest represents the mean of the posterior distribution, and the stability of the estimate, on which inferences are made, is known as a credible interval (akin to conventional confidence intervals), which comprises the 2.5th and 97.5th percentiles from the posterior distribution (Ozechowski, 2014).

Parameters in these models are estimated using Markov Chain Monte Carlo (MCMC) estimation. MCMC is an iterative process in which a prior distribution is specified and posterior values for each parameter are estimated over many iterations, which in turn are used to construct the posterior distribution. MCMC is initiated from at least two randomly selected starting points to facilitate convergence of the iteration process (Zyphur & Oswald, 2013). Convergence is indicated graphically as well as statistically via the potential scale reduction (PSR; Asparouhov & Muthén, 2010). PSR indexes the ration of total variance across chains to the pooled variance within a chain. Smaller PSR values (e.g., PSR < 1.05) indicate that convergence has occurred. Because conventional model fit indices are based on maximum-likelihood estimation, Mplus provides an alternative fit index to evaluate model fit, the posterior predictive *p* (PPP) value, which is less sensitive than chi-square testing to model misspecification. A PPP value greater than .05 indicates good model fit (Asparouhov & Muthén, 2010). Specifically, Muthén and Asparouhov (2012) advocated using prior distributions with small variances for cross-loading factors and residual covariances. Following their recommendations, we specified prior distributions for these parameters of a mean of 0 and variance of .01, which translated to standardized values ranging between –0.2 and .2. In addition to the advantages just discussed, Bayesian estimation has the well-documented advantage over CFA of yielding optimal estimates, and more consistent convergence, with small samples (van de Schoot et al., 2014). Thus, Stage 4 of the analysis consisted of Bayesian structural equation modeling (BSEM) for the four-factor model that failed to converge with CFA specification. We specified this model as a three-level model with sessions nested within individuals within therapists. We estimated models with informative priors using factor loadings and standard errors estimated in the EFA (see Table 1) for the factor loadings in the BSEM, and small variance priors (.01) for cross-loading items and residual covariances.

RESULTS

Stage 1. Item Selection

One-way random ICC(1, 2) was calculated for each item on each of the three coding scales within each of the three sample pools to examine interrater reliability. ICCs were interpreted based on Cicchetti's (1994) criteria for classifying ICC magnitudes: below .40 is poor, .40 to .59 is fair, .60 to .74 is good, and .75 to 1.00 is excellent. As described in the Method section, 13 items did not meet inclusion criteria (ICC > .30 and $p < .10$ OR skew and kurtosis < 10 in all three samples). ICCs described here are for the remaining 38 items: 11 MDFT, 18 BSFT, nine FFT.

ICCs for the 11 MDFT items were mostly in the good-to-excellent range for the MDFT sample, ranging from .64 to .90. Two items had ICCs below this range at .58 and .38. ICCs for MDFT items were lower in the BSFT sample: Four ranged from .62 to .71, five ranged from .41 to .58, one was .33, and two were below .30 but were retained because significance was less than $p = .10$. In the FFT sample, MDFT items ranged from .79 to .90. On the BSFT rating scale, the majority of items had good or excellent ICCs in the MDFT sample ranging from .62 to .93, whereas three items ranged from .51 to .58. In the BSFT sample, six items ranged from .60 to .78, six items ranged from .40 to .56, four items ranged from .30 to .39, and two were below .30 but retained because significance was less than $p = .10$. In the FFT sample, most BSFT items ranged from .61 to .83, with two in the fair range (.51 and .52). For the FFT scale, nearly all items had good or excellent ICCs in the MDFT sample, ranging from .63 to .80. One item was .31 and one was .30; both were retained due to $p < .10$. In the BSFT sample, five items ranged from .62 to .70, and four ranged from .45 to .59.items. In the FFT sample, three items ranged from .67 to .71, three ranged from .54 to .59, and one was .36.

Stage 2. Exploratory Factor Analysis

EFA was conducted on half of the sample to determine the optimal factor structure. All 38 items meeting ICC inclusion criteria were included in EFA and solutions ranging from one to six factors were extracted. In the original model including all 38 items, eigenvalues were 9.67 for one factor, 4.16 for two factors, 2.72 for three factors, 2.18 for four factors, 1.73 for five factors, and 1.62 for six factors. Factor solutions for four, five, and six factors were examined, and the four-factor solution was deemed most interpretable. Ten items that did not load above .40 on any factor were trimmed from the model (five MDFT, three BSFT, two FFT). Three additional BSFT items that consistently loaded on a separate Parent Engagement factor but interfered with model fit and interpretation were trimmed as well. The final four-factor solution included 25 items: six MDFT, 12 BSFT, seven FFT. Table 1 displays the items organized by loadings on derived factors. Modest correlations between factors indicated substantial factor differentiation (nonoverlap): Factor 1 and Factor 2: $r = .33$, Factor 1 and Factor 3: $r = .14$, Factor 1 and Factor 4: $r = .01$, Factor 2 and Factor 3: $r = .13$, Factor 2 and Factor 4: $r = .15$, Factor 3 and Factor 4: $r = .16$.

Factor 1 was named *Interactional Change* and included 10 items: one MDFT, seven BSFT, two FFT. Three BSFT items and one FFT item were dropped from analysis at the CFA stage (see next) and thus were not included on the final version of this factor. Factor loadings for the six items on the final version of Factor 1 ranged from .98 (BSFT: Stimulates dialogues/Directs enactments) to .55 (BSFT:

TABLE 1
Results of Exploratory Factor Analysis on Reliably Coded Fidelity Items for Three Family Therapy Models

	Family Therapy Model	*Factor 1: Interactional Change*	*Factor 2: Relational Reframe*	*Factor 3: Adolescent Engagement*	*Factor 4: Relational Emphasis*
Factor Eigenvalue		8.05	3.25	2.27	1.89
Factor 1: Interactional Change					
Stimulates Dialogues/Directs Enactments	BSFT	**.984**	−.124	−.029	−.003
Coaches Interactions in Session	TBRS	**.940**	−.027	−.001	.003
Moves Close to Direct Restructuring Maneuvers	BSFT	**.938**	−.052	−.031	−.045
Remains Decentralized After Enactment Emerges	BSFT	**.782**	.131	.012	−.139
Conducts In-Session Exercises for New Behaviors	FFT	**.586**	.363	.062	−.051
Focuses on Present Interactions	BSFT	**.551**	.319	.151	−.002
Actively Directs and Elicits New Behaviors[a]	BSFT	**.530**	.147	−.004	.113
Follows Nonverbals to Elicit Participation[a]	BSFT	**.469**	.022	.073	.119
Intensifies/Highlights Interactions[a]	BSFT	**.486**	.208	−.025	.141
Manages or Responds to Negativity[a]	FFT	**.449**	.352	−.056	.032
Factor 2: Relational Reframe					
Provides a Family-Focused Rationale for Change	FFT	-.075	**.855**	−.075	.116
Offers More Positive View of Problems/Family	BSFT	-.003	**.703**	.286	−.060
Uses Meaning-Change Interventions (Reframe)	FFT	.023	**.695**	−.023	.258
Maintains a Relational Focus	FFT	.219	**.509**	.020	.424
Provides Information/Guidance for New Skill	FFT	.462	**.505**	.029	−.059
Develops Relational Reframe for Youth Problems	TBRS	-.220	**.489**	.051	−.095
Targets Adult Participants for Change	TBRS	.157	**.450**	−.272	.005
Factor 3: Adolescent Engagement					
Joins With Children/Adolescents	BSFT	-.013	−.018	**.862**	.066
Targets Interventions Toward Youth	BSFT	.099	.196	**.823**	.038
Explores Adolescent Ecosystem	TBRS	-.399	−.158	**.446**	−.074
Supports Adolescent Investment in Therapy	TBRS	-.116	.128	**.421**	.029
Factor 4: Relational Emphasis					
Gathers Information on Relationship Functions	FFT	-.129	.016	−.126	**.865**
Asks Clarifying Questions/Focuses on Process	BSFT	-.219	.016	.124	**.602**
Enhances Family Attachment/ Communication	TBRS	.337	−.094	.085	**.487**
Connects With All Family Members	BSFT	.062	−.008	.270	**.470**

Note: BSFT = brief strategic family therapy; TBRS = Therapist Behavior Rating Scale ; FFT = functional family therapy.
[a]This item was eliminated at the confirmatory factor analysis phase to achieve adequate model fit.
Bold font indicates the primary factor onto which the given item loads.

Focuses on present interactions). Factor 2 was named *Relational Reframe* and included seven items: two MDFT, one BSFT, four FFT. Factor loadings ranged from .86 (FFT: Provides a family-focused rationale for change) to .45 (MDFT: Targets adult participants for change). Factor 3 was named *Adolescent Engagement* and included four items: two MDFT, two BSFT. Factor loadings were above .80 for both BSFT items (Joins with children/adolescents, Targets interventions toward youth) and were .45 and .42 for MDFT items (Explores adolescent ecosystem, Supports adolescent investment in therapy). Factor 4 was named *Relational Emphasis* and had four items: one MDFT, two

BSFT, one FFT. Factor loadings ranged from .87 (FFT: Gathers information on relationship functions) to .47 (BSFT: Connects with all family members).

Stage 3. Confirmatory Factor Analysis

The four-factor EFA solution was confirmed on the remaining half sample using CFA. First, preliminary CFA models were calculated for each factors individually, and modification indices were used to guide model adjustments until adequate fit was achieved. To achieve fit on the Interactional Change factor, four items were trimmed from the final factor model (see Table 1). Thus, 21 items progressed from the EFA stage to the final confirmed four factors. Model fit for each factor in the CFA model was evaluated using chi-square, RMSEA, and CFI. Fit indices for Interactional Change were $\chi^2(9) = 16.65$, $p = .05$; RMSEA = .08, 90% confidence interval (CI) [.00, .13], CFI = .99. Fit indices for Relational Reframe were $\chi^2(13) = 19.41$, $p = .11$; RMSEA = .06, 90% CI [.00, .11], CFI = .98. Fit indices for Adolescent Engagement were $\chi^2(2) = 4.08$, $p = .23$; RMSEA = .08, 90% CI [.00, .20], CFI = .99. Fit indices for Relational Emphasis were $\chi^2(2) = 2.69$, $p = .26$; RMSEA = .05, 90% CI [.00, .18], CFI = .99. Evaluation of fit indices indicated that model fit was adequate for each of the four factors individually. However, when combined into a single CFA model, the full four-factor model failed to converge. Measurement invariance models attempting to confirm relative equivalence of the four-factor structure across all three sample pools (MDFT, BSFT, FFT) also failed to converge.

Stage 4. Bayesian Structural Equation Modeling

To confirm the four-factor structure, we estimated a three-level BSEM model with small variance priors for cross-loading items and residual covariances. This model converged (PSR = 1.001 replicated over numerous iterations) and showed good fit to the data (*PPP* = .441). Table 2 displays the factor loadings estimated in BSEM. Items loaded robustly on hypothesized factors, and with few exceptions, were minimally correlated with cross-loading factors. Likewise, residual correlations were small ($M = .004$, range = −0.16 to 0.15). Again, the derived factors were modestly correlated, indicting factor differentiation: Factor 1 and Factor 2: $r = .36$, Factor 1 and Factor 3: $r = .18$, Factor 1 and Factor 4: $r = .05$, Factor 2 and Factor 3: $r = .04$, Factor 2 and Factor 4: $r = .20$, Factor 3 and Factor 4: $r = .30$.

DISCUSSION

This empirical distillation of three manualized FT models for adolescent behavior problems yielded four core practice elements, summarized as follows. The first element was labeled Interactional Change (defined by six treatment techniques) and included interventions in which therapists allow or prompt family members to interact with one another naturally to assess family dynamics and direct in-session interactions among members to promote more effective ways of relating. By creating opportunities to communicate more meaningfully, therapists support families in developing new relational skills. The second element, Relational Reframe (seven techniques), involved therapist efforts to transform symptom-focused and/or adolescent-focused perceptions of clinical problems into a new understanding of those problems as being fundamentally relational, thereby motivating families to pursue changes in family relationships as the primary clinical solution. The third element, Adolescent Engagement (four techniques), described interventions in which therapists join with adolescents by seeking their unique points of view and foster treatment engagement by presenting family therapy as an opportunity to address personally meaningful issues within and outside the family. Last, Relational Emphasis (four techniques) contained interventions that focus on the family as a whole, assessing systemic attributions and processes and intervening to improve overall family functioning. Not coincidentally, these four elements map closely onto the framework of the structural family therapy model (see Minuchin & Fishman, 1981), a common progenitor of manualized FT models and for the family therapy approach in general.

We have strong confidence in the clinical and psychometric validity of study findings. The study sample—videotaped therapy sessions—was certified by model developers as being highly faithful to their respective FT models. The distillation tools—model-specific observational fidelity scales—were produced by model developers for the explicit purpose of evaluating the implementation of essential model strategies and techniques. The analytic methods—observational coding and multilevel factor analysis—adhered to rigorous principles of data reliability and factor resolution. All three models contributed at least one treatment technique to three of the four distilled elements (the exception is that Adolescent Engagement does not contain an FFT technique, though model experts confirm that adolescent engagement is core to the model). Thus there is strong justification that the derived factors are indeed core practice elements of the three models in question.

Although hardly exhaustive of the FT approach, the four core elements are a representative foundation of common FT strategies. Certainly, the 21 techniques that collectively define the elements do not capture the full complement of interventions prescribed by the three models. For example, all manualized FT models designed for ABPs (see also Henggeler & Schaeffer, 2016; Slesnick, Erdem, Bartle-Haring, & Brigham, 2013) feature some degree of case management interventions, including home-based wrap-around services—an important category of interventions

TABLE 2
Results of Three-Level Bayesian Structural Equation Modeling on Fidelity Items From Three Family Therapy Models

	Factor 1: Interactional Change	*Factor 2: Relational Reframe*	*Factor 3: Adolescent Engagement*	*Factor 4: Relational Emphasis*
Coaches interactions in session	***.891***	−.016	−.018	−.033
Moves close to direct restructuring maneuvers	***.884***	−.029	−.007	.006
Stimulates dialogues/Directs enactments	***.823***	−.061	−.015	−.038
Remains decentralized after enactment emerges	***.802***	−.056	−.001	−.002
Conducts in-session exercises for new behaviors	***.667***	.115	−.027	.012
Focuses on present interactions	***.625***	.071	.091	.061
Provides a family-focused rationale for change	−.093	***.848***	−.006	−.028
Utilizes meaning-change interventions (reframe)	.002	***.735***	−.064	.006
Offers more positive view of problems/ family	−.046	***.645***	.068	.053
Provides information/guidance for new skill	*.196*	***.529***	.022	−.066
Maintains a relational focus	.025	***.528***	.072	.170
Targets adult participants for change	.011	***.492***	−.131	−.036
Develops relational reframe for youth problems	−.105	***.432***	.040	−.094
Joins with children/adolescents	.005	−.014	***.882***	.043
Targets interventions toward youth	.084	.062	***.748***	−.049
Supports adolescent investment in therapy	−.044	−.016	***.464***	−.022
Explores adolescent ecosystem	−.133	−.060	***.406***	−.013
Gathers information on relationship functions	−.079	−.013	−.083	***.809***
Asks clarifying questions/Focuses on process	−.059	−.007	.015	***.643***
Enhances family attachment/ communication	.080	.044	−.018	***.467***
Connects with all family members	.092	.011	.120	***.465***

Note: Items in bold are hypothesized to load on the corresponding factor. Italicized items fall outside the 95% credibility interval. Bold font indicates the primary factor onto which the given item loads.

that fell outside the purview of the current study. Even so, the 21 distilled techniques are commonly shared FT interventions that are linked to client outcomes for all three observed models and therefore hold great potential for dissemination as effective practices.

Results of this empirical distillation process differed from results of the conceptual process just described (Hogue et al., 2017) in three main ways. First, an entirely new element, Relational Emphasis, emerged. This element makes explicit the foundational concept of the FT approach: Family relations are the primary targets of assessment and change-making interventions. Therapists who work with families but maintain a fundamentally individualistic and/or intrapsychic approach are not operating with this concept as a core value. Second, two elements from the conceptual distillation (Family Behavior Change, Family Restructuring) effectively merged into a single element (Interactional Change). The new element blends FT interventions aimed at changing observable family behaviors with those aimed at changing underlying (i.e., structural-level) family attributions, roles, and relational processes. Both types of change appear advantageous for addressing ABPs. Third, interventions in the conceptual element Family Engagement that pertain to engaging caregivers in treatment were stripped away during empirical distillation, leaving the reduced element Adolescent Engagement in its place. Caregiver-focused engagement interventions instead aligned with the Relational Reframe and Relational Emphasis elements. There is no doubt that all three models emphasize the need to cultivate and balance alliances with multiple family members. This result from empirical distillation could signify that caregiver engagement is conceptualized quite differently across the three models, foreclosing the possibility of a unified element and/or that caregiver engagement

interventions are an embedded feature of relational reframing and relational focus processes, whereas efforts to engage adolescents constitute a distinct therapeutic agenda.

These deviations from the conceptual distillation process highlight the potential value of empirical distillation procedures, which hew more closely to the observable realities of EBI implementation and thereby afford opportunities to enrich or even correct conceptually based results. Indeed, empirical procedures open the door to distillation solutions that might differ meaningfully across service contexts. For example, how might the current FT elements, derived from sessions representing training and research activities under tight control of model developers, diverge if study methods were applied to sessions from community treatment settings wherein model fidelity varies to a greater extent (Hallgren et al., 2018)? Such potential variations will be increasingly relevant as core element EBIs gain traction in community-based practice. Along these lines, the path seems clear to distill core elements of other EBIs for ABPs, principally cognitive-behavioral therapy (Hogue et al., 2018; McCart & Sheidow, 2016).

It is critical to underscore that core elements are not equivalent to manualized treatments. In addition to discrete intervention techniques, treatment manuals invariably articulate principles of treatment coordination—rules for the timing, sequencing, and client- and context-specific targeting of interventions—that constitute the unique parameters and implementation nuances of a given model (Chorpita et al., 2005). Coordination principles determine, for example, how rigidly versus flexibly a therapist should implement model content, as well as the recommended balance between fidelity versus adaptation for individual cases or clinical groups (McHugh, Murray, & Barlow, 2009). Core elements of EBIs thus cannot supplant full treatment models or be utilized effectively as "brief" versions.

Study Strengths and Limitations

Study strengths include the innovative empirical approach to EBI distillation, notably the rigorous observational data collection and analytic methods, and the diversity of the sample, which supports the generalizability of findings. One study limitation was exclusive focus on the extensiveness (i.e., adherence), rather than the expertise (i.e., competence), with which therapists delivered treatment techniques. Therapist expertise in implementing specific techniques is quite difficult to judge reliably (Webb, DeRubeis, & Barber, 2010) and is not directly germane to distilling which techniques are core across models. A second limitation was inability to demonstrate measurement invariance of the four-factor solution, which leaves open the possibility that one of the sample pools exerted a disproportionately strong or weak influence on the distillation process. This possibility is contraindicated to some degree by the representative distribution of intervention techniques from all three models across all four factors. We plan to conduct additional analyses of measurement invariance on this sample in a follow-up study that will leverage item response theory; those complex analytic methods introduce an alternative conceptualization of the sample data that is beyond the scope of the current study.

One important study feature is both strength and limitation. The three sample pools represent diverse characteristics along the continua of ABP populations, treatment phases, and intervention contexts: a prevention trial involving at-risk young adolescents, a training sample of adolescents in community-based care for behavioral health problems, and a set of clinical trials involving older adolescents meeting diagnostic criteria for conduct or substance use disorder. The benefits of selecting a sample that supports generalizability across broad spectrums of clinical problems and treatment contexts are counterbalanced by sacrifice in the acuity with which findings pertain to teenagers with specific clinical disorders at specific junctures in treatment. Also, because of this sample diversity, it was not possible to use identical methods across samples when selecting high-fidelity sessions. Finally, although observing high-fidelity FT sessions was necessary to define core techniques that validly represent the manualized models of interest, it remains unknown whether the distilled elements will be identifiable or viable in routine clinical conditions. To address this gap, the authors are attempting to verify the structure and content of the four FT elements derived in this study within a new pool of ABP cases treated in usual care by community clinicians with varying degrees of FT allegiance and training.

Clinical Implications

Evidence supporting the effectiveness (Chorpita et al., 2017; Weisz et al., 2012), sustainability (Weisz et al., 2018), and perceived value (Chorpita et al., 2015; Southam-Gerow et al., 2014) of core EBIs in routine practice continues to mount. Still, the potency of the FT approach in naturalistic form remains virtually untested (Riedinger et al., 2017). Is it reasonable to believe that nonmanualized interventions governed by core FT elements are a viable alternative to manualized models for treating ABPs in usual care? If the success of FT techniques for ABPs depends fundamentally upon the implementation boost provided by quality assurance procedures such as those mandated by manualized FT models, then core element FT delivered without substantial support by model experts may be ineffective. Of note, one controlled trial evaluated nonmanualized (i.e., naturalistic) FT delivered by community therapists as the routine standard of care, finding that it outperformed naturalistic non-FT services for improving ABPs (Hogue et al., 2015). Moreover, the naturalistic FT condition exceeded a research-defined benchmark for adherence to FT techniques and equaled a benchmark for

long-term outcomes (Hogue, Dauber, & Henderson, 2017). If replicated and expanded, such results might influence providers to weigh the feasibility of cultivating core element FT services—perhaps factoring in upgrades to local FT supervision and quality procedures (Hogue et al., 2013)—against the formidable barriers to importing a manualized FT model.

Of course this begs the question of what core element FT services might look like. That is, how can distilled FT techniques be implemented effectively in routine care? There are at least three implementation approaches with potential merit. First, core FT elements can be translated into a standardized quality assurance system—a de facto "manual" and companion fidelity procedures—that facilitates decision making about when and with whom to use which combination of elements. This has been accomplished with core EBI elements for other youth disorders (e.g., Chorpita & Weisz, 2009). Quality assurance systems for core EBIs have proven to equal or surpass conventional manualized treatment as well as usual care for youth behavior problems (e.g., Chorpita et al., 2017; Weisz et al., 2012). Second, core FT elements can be documented in a flexible clinical protocol to be used independently by community clinicians without standardized training and fidelity procedures. Therapists would be invited to invoke their own conceptualizations about the appropriate timing, sequence, and relative emphasis of each FT element for each case, adjusting FT delivery based on client responsiveness (see Lyon & Koerner, 2016). Third, core FT techniques can be converted to a utilitarian adherence scale that provides empirically derived FT implementation guideposts for line clinicians treating ABPs (see Stirman et al., 2018, for an example involving cognitive-behavioral therapy). Although therapists are notoriously unreliable in reporting on their own adherence to virtually every variety of EBI (e.g., Hurlburt, Garland, Nguyen, & Brookman-Frazee, 2010), FT for ABPs is one arena in which therapist self-report of treatment adherence has shown reasonable concordance with observer ratings (Chapman, McCart, Letourneau, & Sheidow, 2013; Hogue et al., 2015). Effective self-monitoring of adherence to FT elements via pragmatic therapist-report tools, in the context of ongoing case planning and supervision, might confer enormous benefits to increasing the amount, and perhaps quality, of nonmanualized FT delivered in ABP services. To advance such efforts we are currently developing online training procedures to increase therapist delivery of core FT techniques for ABPs via video-based training in accurate FT self-monitoring combined with tailored feedback on self-reported FT use (National Institute on Drug Abuse: R34DA044740). As core elements distillation continues to progress for numerous EBIs, the field will discover the full potential and limits of its utility in various treatment contexts and/or with specific presenting problems.

ACKNOWLEDGMENTS

We acknowledge the dedicated work of the observational coders for this project (in addition to author MB): Kate Feiner, Alma Hidalgo, Tiffany John, Naomi Krohner, Tessa Long, Alexis Nager, Nkem Osian, Nicole Piazza, Alana Ribowsky, Lauren Ryan, Maddison Schiafo, Monique Sledd, and Elise Yenne. We also acknowledge the contributions of the Expert Review Panel to conceptualizing and interpreting study data: Guy Diamond, Genoveva Garcia, Olga Hervis, Silvia Kaminsky, Annie Niermann, and Michael Robbins. We are extremely grateful to Ms. Hervis and Ms. Kaminsky (Family Therapy Training Institute of Miami) for providing session recordings of Brief Strategic Family Therapy and for training coders; and to Holly Waldron (Oregon Research Institute) for providing recordings of Functional Family Therapy.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

FUNDING

Preparation of this article was supported by the National Institute on Drug Abuse (R01DA037496; PI: Hogue).

REFERENCES

Alexander, J., Waldron, H., Robbins, M., & Neeb, A. (2013). *Functional Family Therapy for Adolescent Behavior Problems*. Washington, DC: American Psychological Association.

Alexander, J. F., Waldron, H. B., Barton, C., & Mas, C. H. (1989). The minimizing of blaming attributions and behaviors in delinquent families. *Journal of Consulting and Clinical Psychology*, *57*, 19–24.

Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*, 411–434. doi:10.1207/s15328007sem1203_4

Asparouhov, T., & Muthén, B. O. (2010). *Bayesian analysis using Mplus: Technical implementation*. Retrived from http://www.statmodel.com/download/BayesAdvantages18.pdf.

Baldwin, S. A., Christian, S., Berkeljon, A., & Shadish, W. R. (2012). The effects of family therapies for adolescent delinquency and substance abuse: A meta-analysis. *Journal of Marital and Family Therapy*, *38*, 281–304. doi:10.1111/j.1752-0606.2011.00248.x

Browne, M., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. Bollen & J. Long (Eds.), *Testing structural models* (pp. 136–159). Newbury Park, CA: Sage Publications.

Chapman, J. E., McCart, M. R., Letourneau, E. J., & Sheidow, A. J. (2013). Comparison of youth, caregiver, therapist, trained, and treatment expert raters of therapist adherence to a substance abuse treatment protocol. *Journal of Consulting and Clinical Psychology*, *81*, 674–680. doi:10.1037/a0033021

Chorpita, B. F., Becker, K. D., & Daleiden, E. L. (2007). Understanding the common elements of evidence-based practice: Misconceptions and

clinical examples. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*, 647–652. doi:10.1097/chi.0b013e318033ff71

Chorpita, B. F., Daleiden, E., Park, A., Ward, A., Levy, M., Cromley, T., … Krull, J. (2017). Child STEPs in California: A cluster randomized effectiveness trial comparing modular treatment with community implemented treatment for youth with anxiety, depression, conduct problems, or traumatic stress. *Journal of Consulting & Clinical Psychology*, *85*, 13–25. doi:10.1037/ccp0000133

Chorpita, B. F., Daleiden, E., & Weisz, J. (2005). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental Health Services Research*, *7*, 5–20.

Chorpita, B. F., & Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: Application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting & Clinical Psychology*, *77*, 566–579. doi:10.1037/a0014565

Chorpita, B. F., Park, A., Tsai, K., Korathu-Larson, P., Higa-McMillan, C. K., Nakamura, B. J., … Krull, J. (2015). Balancing effectiveness with responsiveness: Therapist satisfaction across different treatment designs in the Child STEPs randomized effectiveness trial. *Journal of Consulting and Clinical Psychology*, *83*, 709–718. doi:10.1037/a0039301

Chorpita, B. F., & Weisz, J. R. (2009). Modular approach to therapy for children with anxiety, depression, trauma, or conduct problems (MATCH-ADTC). *Proprietary material*. Retrived from http://www.practicewise.com/portals/0/MATCH_public/index.html.

Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*, 284–290. doi:10.1037/1040-3590.6.4.284

Fleiss, J. (1981). Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement*, *5*, 105–112. doi:10.1177/014662168100500115

Garland, A. F., Bickman, L., & Chorpita, B. F. (2010). Change what? Identifying quality improvement targets by investigating usual mental health care. *Administration and Policy in Mental Health and Mental Health Services Research*, *37*, 15–26. doi:10.1007/s10488-010-0279-y

Garland, A. F., Hawley, K. M., Brookman-Frazee, L. I., & Hurlburt, M. (2008). Identifying common, core elements of evidence-based practice for children with disruptive behavior disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*, 505–514. doi:10.1097/CHI.0b013e31816765c2

Hallgren, K., Dembe, A., Pace, B., Imel, Z., Lee, C., & Atkins, D. (2018). Variability in motivational interviewing adherence across sessions, providers, sites, and research contexts. *Journal of Substance Abuse Treatment*, *84*, 30–41. doi:10.1016/j.jsat.2017.10.011

Henggeler, S. W., & Schaeffer, C. M. (2016). Multisystemic therapy: Clinical overview, outcomes, and implementation research. *Family Process*, *55*, 514–528. doi:10.1111/famp.12232

Hervis, O. E., & Robbins, M. (2015). *Brief Strategic Family Therapy Fidelity Rating Scale*. Miami, FL: Family Therapy Training Institute of Miami.

Hogue, A., Bobek, M., Dauber, S., Henderson, C. E., McLeod, B. D., & Southam-Gerow, M. A. (2017). Distilling the core elements of family therapy for adolescent substance use: Conceptual and empirical solutions. *Journal of Child & Adolescent Substance Abuse*, *26*, 437–453. doi:10.1080/1067828X.2017.1322020

Hogue, A., Dauber, S., Henderson, C., Bobek, M., Johnson, C., Lichvar, E., & Morgenstern, J. (2015). Randomized trial of family therapy versus non-family treatment for adolescent behavior problems in usual care. *Journal of Clinical Child and Adolescent Psychology*, *44*, 954–969. doi:10.1080/15374416.2014.963857

Hogue, A., Dauber, S., & Henderson, C. E. (2017). Benchmarking family therapy for adolescent behavior problems in usual care: Fidelity, outcomes, and therapist performance differences. *Administration and Policy in Mental Health and Mental Health Services Research*, *44*, 626–641. doi:10.1007/s10488-016-0769-7

Hogue, A., Dauber, S., Lichvar, E., Bobek, M., & Henderson, C. (2015). Validity of therapist self-report ratings of fidelity to evidence-based practices for adolescent behavior problems: Correspondence between therapists and observers. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*, 229–243. doi:10.1007/s10488-014-0548-2

Hogue, A., Dauber, S., Samuolis, J., & Liddle, H. (2006). Treatment techniques and outcomes in multidimensional family therapy for adolescent behavior problems. *Journal of Family Psychology*, *20*, 535–543. doi:10.1037/0893-3200.20.4.535

Hogue, A., Henderson, C. E., Becker, S. J., & Knight, D. K. (2018). Evidence base on outpatient behavioral treatments for adolescent substance use, 2014-2017: Outcomes, treatment delivery, and promising horizons. *Journal of Clinical Child and Adolescent Psychology*, *47*, 499–526. doi:10.1080/15374416.2018.1466307

Hogue, A., Henderson, C. E., Dauber, S., Barajas, P., Fried, A., & Liddle, H. (2008). Treatment adherence, competence, and outcome in individual and family therapy for adolescent behavior problems. *Journal of Consulting & Clinical Psychology*, *76*, 544–555. doi:10.1037/0022-006X.76.4.544

Hogue, A., Liddle, H., Becker, D., & Johnson-Leckrone, J. (2002). Family-based prevention counseling for high-risk young adolescents: Immediate outcomes. *Journal of Community Psychology*, *30*, 1–22. doi:10.1002/(ISSN)1520-6629

Hogue, A., Liddle, H., Rowe, C., Turner, R., Dakof, G. A., & LaPann, K. (1998). Treatment adherence and differentiation in individual versus family therapy for adolescent substance abuse. *Journal of Counseling Psychology*, *45*, 104–114. doi:10.1037/0022-0167.45.1.104

Hogue, A., Liddle, H. A., & Rowe, C. (1996). Treatment adherence process research in family therapy: A rationale and some practical guidelines. *Psychotherapy: Theory, Research, Practice, and Training*, *33*, 332–345. doi:10.1037/0033-3204.33.2.332

Hogue, A., Liddle, H. A., Singer, A., & Leckrone, J. (2005). Intervention fidelity in family-based prevention counseling for adolescent problem behaviors. *Journal of Community Psychology*, *33*, 191–211. doi:10.1002/(ISSN)1520-6629

Hogue, A., Ozechowski, T., Robbins, M., & Waldron, H. (2013). Making fidelity an intramural game: Localizing quality assurance procedures to promote sustainability of evidence-based practices in usual care. *Clinical Psychology: Science and Practice*, *20*, 60–77.

Hurlburt, M. S., Garland, A. F., Nguyen, K., & Brookman-Frazee, L. (2010). Child and family therapy process: Concordance of therapist and observational perspectives. *Administration and Policy in Mental Health and Mental Health Services Research*, *37*, 230–244. doi:10.1007/s10488-009-0251-x

Lyon, A., & Koerner, K. (2016). User-centered design for psychosocial intervention development and implementation. *Clinical Psychology: Science & Practice*, *23*, 180–200.

McCart, M. R., & Sheidow, A. J. (2016). Evidence-based psychosocial treatments for adolescents with disruptive behavior. *Journal of Clinical Child & Adolescent Psychology*, *45*, 529–563. doi:10.1080/15374416.2016.1146990

McDonald, R., & Ho, M. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, *7*, 64–82.

McHugh, R., Murray, H., & Barlow, D. H. (2009). Balancing fidelity and adaptation in the dissemination of empirically supported treatments: The promise of transdiagnostic interventions. *Behavior Research and Therapy*, *47*, 946–953. doi:10.1016/j.brat.2009.07.005

Minuchin, S., & Fishman, H. C. (1981). *Family therapy techniques*. Cambridge, MA: Harvard.

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, *17*, 313–335. doi:10.1037/a0026802

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (Seventh ed ed.). Los Angeles, CA: Muthen & Muthen.

Ozechowski, T. J. (2014). Empirical Bayes MCMC estimation for modeling treatment processes, mechanisms of change, and clinical outcomes in small samples. *Journal of Consulting and Clinical Psychology*, *82*, 854–867. doi:10.1037/a0035889

Ozechowski, T. J., & Waldron, H. (2016). *Functional Family Therapy Therapist Adherence Rating Scale*. Eugene, OR: Oregon Research Institute.

Riedinger, V., Pinquart, M., & Teubert, D. (2017). Effects of systemic therapy on mental health of children and adolescents: A meta-analysis. *Journal of Clinical Child and Adolescent Psychology*, *46*, 880–894. doi:10.1080/15374416.2015.1063427

Robbins, M., Alexander, J., & Turner, C. (2000). Disrupting defensive interactions in family therapy with delinquent adolescents. *Journal of Family Psychology*, *14*, 688–701.

Robbins, M., Feaster, D., Horigian, V., Puccinelli, M., Henderson, C. E., & Szapocznik, J. (2011). Therapist adherence in brief strategic family therapy for adolescent drug abusers. *Journal of Consulting & Clinical Psychology*, *79*, 43–53. doi:10.1037/a0022146

Robbins, M. S., Alexander, J. F., Newell, R. M., & Turner, C. W. (1996). The immediate effect of reframing on client attitude in family therapy. *Journal of Family Psychology*, *10*, 28–34. doi:10.1037/0893-3200.10.1.28

Rohde, P., Waldron, H., Turner, C., Brody, J., & Jorgensen, J. (2014). Sequenced versus coordinated treatment for adolescents with comorbid symptoms and substance use disorders. *Journal of Consulting & Clinical Psychology*, *82*, 342–348. doi:10.1037/a0035808

Shrout, P., & Fleiss, J. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420–428.

Slesnick, N., Erdem, G., Bartle-Haring, S., & Brigham, G. S. (2013). Intervention with substance-abusing runaway adolescents and their families: Results of a randomized clinical trial. *Journal of Consulting & Clinical Psychology*, *81*, 600–614. doi:10.1037/a0033463

Southam-Gerow, M. A., Daleiden, E. L., Chorpita, B. F., Bae, C., Mitchell, C., Faye, M., & Alba, M. (2014). MAPping Los Angeles county: Taking an evidence-informed model of mental health care to scale. *Journal of Clinical Child & Adolescent Psychology*, *43*, 190–200. doi:10.1080/15374416.2013.833098

Stirman, S. W., Marques, L., Creed, T. A., Gutner, C. A., DeRubeis, R., Barnett, P. G., … La Bash, H. (2018). Leveraging routine clinical materials and mobile technology to assess CBT fidelity: The Innovative Methods to Assess Psychotherapy Practices (imAPP) study. *Implementation Science*, *13*, 69. doi:10.1186/s13012-018-0779-9

Substance Abuse and Mental Health Services Administration. (2012). *National Registry of Evidence-based Programs and Practices (NREPP). Intervention summaries: Multisystemic therapy for juvenile offenders, brief strategic family therapy, and multidimensional family therapy*. Rockville, MD: Govt Printing Office (SAMHSA).

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston, MA: Allyn & Bacon.

Tanner-Smith, E. E., Wilson, S. J., & Lipsey, M. W. (2013). The comparative effectiveness of outpatient treatment for adolescent substance abuse: A meta-analysis. *Journal of Substance Abuse Treatment*, *44*, 145–158. doi:10.1016/j.jsat.2012.05.006

van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & van Aken, M. A. G. (2014). A gentle introduction to bayesian analysis: Applications to developmental research. *Child Development*, *85*, 842–860. doi:10.1111/cdev.12169

Waldron, H., Slesnick, N., Brody, J., & Turner, C. W. (2001). Four- and seven-month treatment outcomes for adolescent substance abuse. *Journal of Consulting & Clinical Psychology*, *62*, 802–813. doi:10.1037/0022-006X.69.5.802

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, *78*, 200–211. doi:10.1037/a0018912

Weisz, J. R., Bearman, S., Santucci, L., & Jensen-Doss, A. (2017). Initial test of a principle-guided approach to transdiagnostic psychotherapy with children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, *46*, 44–58. doi:10.1080/15374416.2016.1163708

Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., … Gibbons, R. D. (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. *Archives of General Psychiatry*, *69*, 274–282. doi:10.1001/archgenpsychiatry.2011.147

Weisz, J. R., Ugueto, A. M., Herren, J., Marchette, L. K., Bearman, S. K., Lee, E. H., … Jensen-Doss, A. (2018). When the torch is passed, does the flame still burn? Testing a "train the supervisor" model for the Child STEPs treatment program. *Journal of Consulting and Clinical Psychology*, *86*, 726–737. doi:10.1037/ccp0000331

Zyphur, M. J., & Oswald, F. L. (2013). Bayesian estimation and inference: A user's guide. *Journal of Management*, *41*, 390–420. doi:10.1177/0149206313501200